

日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT



別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1999年 9月16日

出願番号
Application Number:

平成11年特許願第261558号

出願人
Applicant(s):

株式会社日立製作所

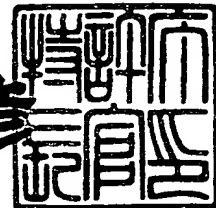
U.S. Appln. Filed 9-12-00
Inventor. K. Iwatsuki et al
Mathingly. Stanger et al
Docket H-942

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年 6月16日

特許庁長官
Commissioner,
Patent Office

近藤 隆彦



出証番号 出証特2000-3045668

【書類名】	特許願
【整理番号】	K99055281
【提出日】	平成11年 9月16日
【あて先】	特許庁長官殿
【国際特許分類】	H04L 12/00
【請求項の数】	32
【発明者】	
【住所又は居所】	神奈川県秦野市堀山下 1 番地 株式会社日立製作所 エ ンタープライズサーバ事業部内
【氏名】	岩月 和子
【発明者】	
【住所又は居所】	神奈川県秦野市堀山下 1 番地 株式会社日立製作所 エ ンタープライズサーバ事業部内
【氏名】	綿貫 達哉
【発明者】	
【住所又は居所】	神奈川県秦野市堀山下 1 番地 株式会社日立製作所 エ ンタープライズサーバ事業部内
【氏名】	宮本 貴久
【発明者】	
【住所又は居所】	神奈川県秦野市堀山下 1 番地 株式会社日立製作所 エ ンタープライズサーバ事業部内
【氏名】	安江 利一
【特許出願人】	
【識別番号】	000005108
【氏名又は名称】	株式会社日立製作所
【代理人】	
【識別番号】	100075096
【弁理士】	
【氏名又は名称】	作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 ネットワーク装置

【特許請求の範囲】

【請求項 1】

他の装置と全二重リンクを介して接続されるネットワーク装置において、
前記他の装置からのフレームの送信を中断させる制御フレームを、前記他の装置
に対して送信するよう送信指示を発行する帯域制御手段と、
前記送信指示に従い前記制御フレームを送信する送信手段とを有し、前記他の装
置から送信されるフレームの送信帯域量を、予め設定された設定帯域量以内に制
御することを特徴とするネットワーク装置。

【請求項 2】

請求項 1 に記載のネットワーク装置において、
前記制御フレームの送信される送信間隔を計数するタイマとを有し、
前記タイマが一定時間を計測する毎に、前記帯域制御手段は、前記制御フレーム
の送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 3】

請求項 1 に記載のネットワーク装置において、
前記帯域制御手段は、前記他の装置から送信されるフレームの送信帯域量が、前
記設定帯域量を超えたときだけ、前記制御フレームの送信指示を発行するもので
あることを特徴とするネットワーク装置。

【請求項 4】

複数の装置の各々に対応して設けられるポートを有し、前記ポートの各々に接
続される全二重リンクを介して前記複数の装置の各々と接続されるネットワーク
装置であって、
前記複数の装置からのフレームの送信を中断させる P A U S E フレームを、前記
ポートから送信させるよう指示する帯域制御手段と、
前記 P A U S E フレームを送信する送信手段とを有し、前記複数の装置から送信
されるフレームの送信帯域量を、予め設定された設定帯域量以内に制御すること
を特徴とするネットワーク装置。

【請求項 5】

請求項 4 に記載のネットワーク装置において、
前記帯域制御手段は、定期的に前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 6】

請求項 4 に記載のネットワーク装置において、
前記設定帯域量は、前記複数の装置から送信されるフレームの送信帯域量の総和に対して設定されており、
前記帯域制御手段は、前記複数の装置からのフレームの送信される送信帯域量の総和が、前記設定帯域量を超えたときだけ、前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 7】

請求項 4 に記載のネットワーク装置において、
前記設定帯域量は、前記複数の装置の各々に対して設定されており、
前記帯域制御手段は、前記複数の装置からのフレームの送信される送信帯域量の各々について前記設定帯域量を超えたときだけ、前記全二重リンクのうちの前記設定帯域量を超えた全二重リンクにのみ、前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 8】

請求項 4 乃至 7 に記載のネットワーク装置において、
前記全二重リンクは、I E E E 8 0 2 . 3 a d のリンク・アグリゲーション規格が適用される複数の全二重リンクであることを特徴とするネットワーク装置。

【請求項 9】

請求項 4、5、7 に記載のネットワーク装置において、
前記帯域制御手段は、複数存在する前記ポートの各々に対して設けられるものであることを特徴とするネットワーク装置。

【請求項 1 0】

請求項 4、6 に記載のネットワーク装置において、
前記制御手段は、複数存在する前記ポートの全てに対してただ 1 つ設けられるも

のであることを特徴とするネットワーク装置。

【請求項 11】

複数の装置の各々に対応して設けられる複数のポートを有し、前記複数のポートの各々に接続される複数の全二重リンクを介して前記複数の装置の各々と接続されるネットワーク装置であって、

前記複数の装置からのフレームの送信を中断させる P A U S E フレームを、前記複数のポートから送信させるよう指示する帯域制御手段と、

前記 P A U S E フレームを送信する送信手段とを有し、前記複数の装置から送信されるフレームの送信帯域量を、予め設定された設定帯域量以内に制御することを特徴とするネットワーク装置。

【請求項 12】

請求項 11 に記載のネットワーク装置において、
前記帯域制御手段は、定期的に前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 13】

請求項 11 に記載のネットワーク装置において、
前記設定帯域量は、前記複数の装置から送信されるフレームの送信帯域量の総和に対して設定されており、

前記帯域制御手段は、前記複数の装置から送信されるフレームの送信帯域量の総和が、前記設定帯域量を超えたときだけ、前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 14】

請求項 11 に記載のネットワーク装置において、
前記設定帯域量は、前記複数の装置毎に、前記複数の全二重リンクに送信されるフレームの送信帯域量の総和に対して設定されており、
前記帯域制御手段は、前記複数の装置毎に前記複数の全二重リンクに送信されるフレームの送信帯域量の総和が、前記設定帯域を超えたときだけ、前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 15】

請求項 11 に記載のネットワーク装置において、
前記設定帯域量は、前記複数の装置の各々に対応して設定されており、
前記帯域制御手段は、前記複数の装置から送信されるフレームの送信帯域量の各々について前記設定帯域量を超えたときだけ、前記全二重リンクのうちの前記設定帯域量を超えた全二重リンクにのみ、前記 P A U S E フレームの送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 16】

複数の装置の各々に対応して設けられる複数のポートを有し、前記複数のポートの各々のポートに対して、全二重リンクを介して前記複数のネットワーク装置が有するポートと接続され、I E E E 8 0 2 . 3 a d のリンク・アグリゲーション規格を適用したネットワーク装置であって、
前記複数の装置からのフレームの送信を中断させる P A U S E フレームを送信する送信指示を発行する帯域制御手段と、
I E E E 8 0 2 . 3 x 規格の P A U S E 機能を備え、前記各々のポート毎に対応して設けられ、前記送信指示を受ける M A C 制御部と、
複数の前記 M A C 制御部を集約し、前記送信指示を前記複数の M A C 制御部に配布するリンク・アグリゲーションサブレイヤとを有することを特徴とするネットワーク装置。

【請求項 17】

請求項 16 に記載のネットワーク装置において、
前記帯域制御手段は、前記複数の M A C 制御部毎に設けられ、前記複数の M A C 制御部のうち自帯域制御手段に対応する M A C 制御部に対して、定期的に前記送信指示を発行するものであることを特徴とするネットワーク装置。

【請求項 18】

請求項 16 に記載のネットワーク装置において、
前記全二重リンクの帯域量の総和を集計するバイトカウンタとを有し、
前記帯域制御手段は、前記複数の M A C 制御部に対してただ 1 つ設けられ、前記帯域量の総和が、ネットワーク管理者が設定する帯域量を超える場合に前記送信

指示を発行するものであることを特徴とするネットワーク装置。

【請求項 19】

IEEE 802.3x 規格の PAUSE 機能を備え、フレームの送信の中断を指示する PAUSE フレームを受信する MAC チップセットと、IEEE 802.3x 規格の PAUSE 機能を備え、前記受信した PAUSE フレームにセットされた PAUSE 時間の間サーバ装置自身からのフレームの送信を中断させる MAC 制御部とを有するサーバ装置に対して、全二重リンクを介して接続されるネットワーク装置であって、

前記サーバ装置から送信されるフレームの送信帯域が前記全二重リンクの物理帯域を使用する割合である帯域使用率と、前記サーバ装置からのフレームの送信が中断される期間である PAUSE 時間と、PAUSE フレームの送信される間隔である PAUSE 送信間隔とが設定されるレジスタ群と、

前記 PAUSE 送信間隔がセットされるタイマとを有し、前記サーバ装置に対して、PAUSE フレームを送信する指示である送信指示を発行する帯域制御手段と、

IEEE 802.3x 規格の PAUSE 機能を備え、前記帯域制御手段から前記送信指示を受ける MAC 制御部と、

IEEE 802.3x 規格の PAUSE 機能を備え、前記送信指示に従い、フレームの送信の中断を指示する PAUSE フレームを生成し、前記サーバ装置に対して前記生成した PAUSE フレームを送信する MAC チップセットとを有することを特徴とするネットワーク装置。

【請求項 20】

IEEE 802.3x 規格の PAUSE 機能を備え、フレームの送信の中断を指示する PAUSE フレームを受信する MAC チップセットと、IEEE 802.3x 規格の PAUSE 機能を備え、前記受信した PAUSE フレームにセットされた PAUSE 時間の間サーバ装置自身からのフレームの送信を中断させる MAC 制御部とを有するサーバ装置に対して、全二重リンクを介して接続されるネットワーク装置であって、

前記サーバ装置から送信されるフレームの帯域量を測定するバイトカウンタと、

前記サーバ装置から送信されるフレームの送信帯域が前記全二重リンクの物理帯域を使用する割合である帯域使用率と、前記サーバ装置からのフレームの送信が中断される期間である P A U S E 時間と、前記バイトカウンタの測定値を読み取る時間の間隔であるトラフィック観測時間とが設定されるレジスタ群と、前記トラフィック観測時間がセットされるタイマとを有し、前記バイトカウンタの測定値を利用して前記 P A U S E 時間を計算し、前記測定値が前記帯域使用率に相当する送信帯域値を超えたときだけ、前記サーバ装置に対して、P A U S E フレームの送信を指示する送信指示を発行する帯域制御手段と、

I E E E 8 0 2. 3 x 規格の P A U S E 機能を備え、前記帯域制御手段から前記送信指示を受ける M A C 制御部と、

I E E E 8 0 2. 3 x 規格の P A U S E 機能を備え、前記送信指示に従い、フレームの送信の中断を指示する P A U S E フレームを生成し、前記サーバ装置に対して前記生成した P A U S E フレームを送信する M A C チップセットとを有することを特徴とするネットワーク装置。

【請求項 2 1】

ネットワーク装置に設けられ、前記ネットワーク装置と全二重リンクを介して接続される他の装置からのフレームの送信帯域量を制御する帯域制御機能であって、

I E E E 8 0 2. 3 x で規定される P A U S E 機能を備える M A C 制御部に対する上位のプロトコルである M A C クライアントと、

定期的に送信される制御フレームの送信間隔値が設定されるレジスタ群と

前記レジスタ群に設定された前記送信間隔値を計数するタイマとを有し、

前記タイマにおいて前記送信間隔値を計数した時に、前記 M A C 制御部に対して、前記他の装置宛てに前記制御フレームの送信を指示するものであることを特徴とする帯域制御機能。

【請求項 2 2】

請求項 2 1 に記載の帯域制御機能において、

前記レジスタ群は、前記他の装置から送信されるフレームの帯域が前記全二重リンクの物理帯域を使用する割合である帯域使用率と、

前記他の装置からのフレームの送信を中断させる期間であり、前記制御フレームにセットされる P A U S E 時間とが設定されるものであることを特徴とする帯域制御機能。

【請求項 2 3】

ネットワーク装置と他の装置とが全二重リンクで接続されているネットワークにおいて前記他の装置からの送信帯域量を制御する帯域制御方法であって、前記ネットワーク装置は、前記他の装置に対して、フレームの送信の中断を指示する制御フレームを定期的に送信し、前記制御フレームを受信した前記他の装置は、フレームの送信を中断することを特徴とする帯域制御方法。

【請求項 2 4】

各々 I E E E 8 0 2 . 3 x で規定される P A U S E 機能を備える第 1 のネットワーク装置および第 2 のネットワーク装置の間を全二重リンクを介して接続されるネットワークにおいて、前記 I E E E 8 0 2 . 3 x で規定される P A U S E フレームを用いて第 1 のネットワーク装置および第 2 のネットワーク装置の間の帯域の制御を行うフレーム送受信方法であって、前記第 1 のネットワーク装置は、前記第 2 のネットワーク装置に対して、フレームの送信を中断させる期間である中断期間がセットされ、フレームの送信の中断を指示する前記 P A U S E フレームを送信し、前記 P A U S E フレームを受信した前記第 2 のネットワーク装置は、フレームの送信を中断し、前記第 2 のネットワーク装置は、前記中断期間が経過した後、フレームの送信を開始することを特徴とするフレーム送受信方法。

【請求項 2 5】

請求項 2 4 に記載のフレーム送受信方法において、前記第 1 のネットワーク装置からの前記 P A U S E フレームの送信は、定期的に行われることを特徴とするフレーム送受信方法。

【請求項 2 6】

請求項 2 4、2 5 に記載のフレーム送受信方法において、

前記第2のネットワーク装置は、前記中断期間中に、前記第1のネットワーク装置から他のPAUSEフレームを受信した場合は、前記他のPAUSEフレームを受信した時点で、前記他のPAUSEフレームにセットされ、前記第2のネットワーク装置からのフレームの送信を中断させる期間である中断期間を開始させることを特徴とするフレーム送受信方法。

【請求項27】

請求項24に記載のフレーム送受信方法において、
前記第1のネットワーク装置からの前記PAUSEフレームの送信は、前記第2のネットワーク装置から送信されるフレームの送信帯域量が、前記全二重リンクの有する物理帯域に対して予め設定された帯域使用量を超えた場合に、実行されることを特徴とするフレーム送受信方法。

【請求項28】

他のネットワーク装置と全二重リンクを用いて接続され、IEEE802.3xで規定されるPAUSE機能を備えるMAC制御部およびMACチップセットを有するネットワーク装置の帯域制御方法であって、
前記他のネットワーク装置から送信されるフレームの送信帯域量を制御する帯域制御手段が、前記他のネットワーク装置に対して、フレームの送信の中断を指示する制御フレームを作成し、
前記帯域制御手段は、前記制御フレームを、メモリに保管させ、
定期的に送信される前記制御フレームの送信間隔が経過した時に、前記帯域制御手段が前記MAC制御部に対して、前記サーバ装置への前記制御フレームの送信を指示し、
前記制御フレームの送信指示を受けたMAC制御部は、前記MACチップセットに対して、前記メモリに対するダイナミック・メモリ・アクセスを指示し、
前記ダイナミック・メモリ・アクセス指示を受けた前記MACチップセットは、前記制御フレームの送信を実行することを特徴とする帯域制御方法。

【請求項29】

IEEE802.3xで規定されるPAUSE機能を備える他のネットワーク装置と全二重リンクを介して接続され、IEEE802.3xで規定されるPA

USE機能を備えるネットワーク装置の帯域制御方法において、
 前記他のネットワーク装置からフレームが送信される送信帯域量を測定し、
 前記送信帯域量の測定値が許容される送信帯域量を超える場合、前記他のネットワーク装置からのフレームの送信が中断される時間であるPAUSE時間を計算し、
 前記PAUSE時間がセットされ、前記他のネットワーク装置に対してフレームの送信の中断を指示するPAUSEフレームを作成し、
 前記PAUSEフレームを送信することを特徴とするネットワーク装置の帯域制御方法。

【請求項 3 0】

請求項 2 9 に記載のネットワーク装置の帯域制御方法において、
 前記送信帯域の測定は、ネットワーク管理者により設定された時間毎に行われることを特徴とするネットワーク装置の帯域制御方法。

【請求項 3 1】

第 1 の装置と、前記第 1 の装置と全二重リンクを介して接続され、前記第 1 の装置に対してフレームの送信を中断させる制御フレームを送信する第 2 の装置との間のフレーム伝送方法において、
 一定の期間毎に前記全二重リンクにおけるフレームの伝送が中断されることを特徴とするフレーム伝送方法。

【請求項 3 2】

第 1 の装置と、前記第 1 の装置と全二重リンクを介して接続され、前記第 1 の装置に対してフレームの送信を中断させる制御フレームを送信する第 2 の装置との間のフレーム伝送方法において、
 前記第 1 の装置が送信したフレームの前記全二重リンクにおける帯域量が予め設定された帯域量を超えた場合に、前記予め設定された帯域を超えた帯域量に対応した期間だけ、前記全二重リンクにおけるフレームの伝送が中断されることを特徴とするフレーム伝送方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、サーバクライアントシステムを構成するネットワーク装置に係り、特に、1台又は複数のサーバからなるサーバファームと、クライアント装置を束ねるLANとの間に置かれてサーバからクライアントへの通信帯域（トラフィック量）制御に関する。

【0002】

【従来技術】

IEEE 802. 3では、ローカル・エリア・ネットワーク（LAN）について規定している。多くのネットワークシステムが、IEEE 802. 3で規定のLANで構成されている。

【0003】

近年、このネットワークシステムの構成装置の主体が、共有バスネットワークを構築するハブ装置から、スイッチングネットワークを構築するLANスイッチへと変化している。この変化に従い、ネットワークシステムを構成する装置のネットワーク・インタフェース同士を接続するリンクも、変化してきている。すなわち、従来は、共有バスで使用される半二重リンクであったものが、全二重リンクへと変化してきている。この全二重リンクとは、リンクに接続する装置が共に、同時に送受信可能なリンクである。また、ネットワークシステムを構成する装置とは、例えば、LANスイッチ等のネットワーク装置、サーバ/PC等のホスト装置等である。

【0004】

全二重リンクは、送信用と受信用にそれぞれ専用の通信線がある。全二重リンクによれば、データ衝突が発生しないので、連続したフレーム送信が可能である。しかし、その送信トラフィックが受信側装置の受信処理能力を超えてしまうとフレーム廃棄が発生する。このフレーム廃棄を回避する手段として、IEEE 802. 3xと呼ばれる標準仕様が規定されている。IEEE 802. 3xは、全二重リンクで接続する装置間のトラフィックを抑制する技術である。

【0005】

IEEE 802.3xでは、全二重リンクで接続される2台の装置において、一方の装置が相手装置からのフレーム送信を一定期間止めたい場合について規定されている。この場合に、相手装置に対して「PAUSE（ポーズ）」と呼ばれる制御フレームを送信する。そのPAUSEフレームを受信した相手装置は、PAUSEフレームで指定される期間、フレームの送信を止めることとなる。

【0006】

具体的な内容を、図15を用いて説明する。図15に示すように、全二重リンク2で接続するネットワーク装置151およびサーバ装置3は、MAC（Media Access Control）チップセット4、5及びMAC制御部6、7を備える。MACチップセット4、5及びMAC制御部6、7がIEEE 802.3x規格のPAUSE機能を備える場合、次のような動作をする。

【0007】

MAC制御部6、7は、上位プログラム等からPAUSE時間の情報を含むPAUSEフレーム送信指示を受ける。MACチップセット4、5は、この送信指示により、指示されたPAUSE時間がセットされたPAUSEフレームを生成する。MACチップセット4、5は、PAUSEフレームを生成した後、PAUSEフレームを送信する。また、MACチップセット4、5は、接続先の装置からPAUSEフレームを受信する。MAC制御部6、7は、受信したPAUSEフレームの内容を解読する。解読の後、MAC制御部6、7は、指定されたPAUSE時間を、内蔵するPAUSEタイマ8、9にセットする。MAC制御部6、7は、セットと同時に、セットした時点での実行中のフレームの送信が終了するのを待つ。MAC制御部6、7は、送信終了後に、PAUSEタイマ8、9をスタートさせる。そして、サーバ装置3は、PAUSE時間が経過するまでの間、フレームの送信を中断する。

【0008】

このPAUSE時間は、1から65535の範囲で指定される。1PAUSE時間の単位は、64バイトのフレーム1個の送信時間である。すなわち、512ビットの送信に費やす時間（512ビット時間）である。

【0009】

一方、近年イントラネットの構成が変化している。この変化により、組織の中心にあるエンタープライズ・サーバ群（サーバ・ファーム）と各部門のクライアント間との通信トラフィックの比率が高くなっている。また、IEEE 802.3z 規格等の高速 LAN や、装置間の高速接続方式であるリンク・アグリゲーション技術（IEEE 802.3ad 規格）が登場してきた。これらの登場により、サーバ装置からの送信トラフィックが大きくなる傾向にある。なお、トラフィックとは、通信帯域量のことであり、送信トラフィックとは、単位時間における送信帯域量のことである。

【0010】

一般に、サーバクライアントシステムでは、末端クライアントが接続する LAN の帯域は小さい。従って、サーバ装置からの送信トラフィックが大きい場合、途中のネットワーク装置で輻輳が発生する。従来のネットワーク装置は、輻輳の発生に伴い、フレームを廃棄していた。

【0011】

【発明が解決しようとする課題】

通常、通信ネットワークシステムを構成する装置においては、その処理能力を超えたトラフィックを受信したときに、フレームを廃棄する。廃棄されたフレームは再送される。従って、さらにトラフィックは増大する。その結果、さらなるフレーム廃棄を招くという悪循環が生じる。

【0012】

このように、フレーム廃棄によるトラフィック制御では、輻輳を解決できない。高トラフィックの送信元は、サーバ・ファームである。従って、サーバ装置自身に送信トラフィックを抑制させる手段が求められる。

【0013】

また、ネットワーク管理者においては、サーバ・トラフィックをアプリケーション単位に帯域制御する。これにより、輻輳を回避しつつネットワーク・ユーザが満足する通信帯域を提供する手段が求められている。

【0014】

上述のように、IEEE 802.3x規格は、2台の装置間で接続先からの送信トラフィックを抑制する目的で提案されている。IEEE 802.3x規格では、PAUSEフレーム送信を起動させて帯域制御を実施する手段は規定されていない。従って、上記の課題解決のためには新たな手段が必要である。

【0015】

本発明の目的は、サーバクライアントシステムを構成するネットワークシステムにおいて、輻輳を回避させることである。また、本発明の目的は、サーバ・トラフィックをネットワーク管理者が設定する単位（アプリケーションや宛先IPアドレス等）で、帯域を制御することである。さらに、本発明の他の目的は、不要な送信遅延を起こすことなく、帯域を制御することである。

【0016】

【課題を解決するための手段】

上記目的を達成するための手段について述べる。ネットワーク装置は、1台または複数のサーバ装置に直接接続して、ネットワークシステムの構成要素となる。ネットワーク装置において、ネットワーク管理者は、各々のサーバ装置からの送信トラフィックに対して、予め任意の帯域を割り当てる。ネットワーク装置は、各々のサーバ装置に対して、定期的にPAUSEフレームを送信するよう送信指示を行う帯域制御手段を有する。ネットワーク装置は、帯域制御手段の送信指示に従い、各々のサーバ装置に対して、全二重リンクで接続された各々のポートから定期的にPAUSEフレームを送信する。PAUSEフレームには、サーバ装置からのフレームの送信を中断させる中断期間がセットされている。サーバ装置は、ネットワーク装置から送信されたPAUSEフレームを受信する。サーバ装置は、受信したPAUSEフレームにセットされた中断期間に関連する期間だけ、ネットワーク装置に対するフレームの送信を中断する。サーバ装置は、中断期間が経過した後に、ネットワーク装置に対するフレームの送信を開始する。このようにサーバ装置のフレームの送信の中断期間を有効に利用することにより、各々のサーバ装置からの送信トラフィックを、予め設定された帯域量以内に制御することが可能となる。

【0017】

また、本発明の他の実施の形態では、上記帯域制御手段は、各々のサーバ装置からの送信トラフィックが予め設定された帯域量を超えたときだけ、PAUSEフレームを送信するようにしてもよい。予め設定された帯域量を超えたときだけPAUSEフレームが送信されることにより、定期的なサーバ装置のフレーム送信の中断による不要な送信遅延を防ぐことが可能となる。

【0018】

【発明の実施の形態】

(実施例1)

以下、本発明の実施例を、図1を用いて説明する。

【0019】

なお、本実施例で記述する時間の単位は、特に指定しない限り、「従来の技術」の欄で説明したIEEE802.3xにおける時間単位(512ビット時間)である。

【0020】

図1は、本発明によるネットワーク装置1と、ネットワーク装置1にリンク2で接続するサーバ装置3を示す。リンク2は、ネットワーク装置1-サーバ装置3の間で、全二重通信を行うことに用いられる物理的な伝送媒体である。ネットワーク装置1及びサーバ装置3は、MACチップセット4、5及びMAC制御部6、7を備える。MACチップセット4、5及びMAC制御部6、7は、共にIEEE802.3x規格のPAUSE機能を備えている。図中、MACチップセット4、5、物理層チップセット、データメモリ14、15は、それぞれハードウェアおよびソフトウェアで構成されている。MAC制御部6、7、帯域制御プログラム10、11、PAUSEタイマ8、9、タイマ12、レジスタ群13は、それぞれソフトウェアで構成されている。

【0021】

本発明による帯域制御プログラム10は、MACクライアントプログラムを含む。帯域制御プログラム10は、タイマ12とレジスタ群13を備える。レジスタ群には、サーバ装置3の送信トラフィックの帯域使用率 n (%)、PAUSE

時間 a 、PAUSE 送信間隔 b を設定するための複数のレジスタが設けられる。
 帯域使用率 n とは、リンク 2 の有する帯域のうち、サーバ装置 3 から送信される
 フレームの送信帯域が、リンク 2 の物理帯域を使用する割合である。PAUSE
 時間 a とは、PAUSE フレームにより、サーバ装置 3 からのフレームの送信が
 止められる時間のことである。PAUSE 送信間隔 b とは、PAUSE フレーム
 の送信される間隔のことである。

【0022】

ネットワーク管理者は、ネットワーク装置 1 とサーバ装置 3 を接続するリンク
 2 の帯域に対しての帯域使用率 n を設定する。帯域使用率 n の値は、レジスタ群
 13 にセットされる。また、PAUSE 送信間隔 b は、予め設定されている。

【0023】

帯域制御プログラム 10 は、PAUSE 送信間隔 b と、帯域使用率 n の値から
 、PAUSE 時間 a を計算する。PAUSE 時間 a の計算の後、帯域制御プログ
 ラム 10 は、PAUSE フレームの送信を指示する。MAC チップセット 4 は、
 この送信指示に従って、指示された PAUSE 時間がセットされた PAUSE フ
 レームを生成する。生成された PAUSE フレームは、データメモリ 14 に保持
 される。また、MAC チップセット 4 は、PAUSE フレームを送信する。

【0024】

PAUSE 時間 a 、PAUSE 送信間隔 b および帯域使用率 n の関係について
 説明する。

【0025】

PAUSE 時間 a 、PAUSE 送信間隔 b および帯域使用率 n には、数 1 の関
 係がある。

【0026】

【数 1】

$$(b - a) / b = n / 100 \quad \cdots \text{(数 1)}$$

即ち PAUSE 機能により、PAUSE 送信間隔 b の時間内でサーバ装置 3 が
 送信可能な時間の比率は $(b - a) / b$ となる。この値がユーザにより設定され
 る $n / 100$ に等しくなるように PAUSE 時間 a を求める。

【0027】

IEEE 802.3x規格では、ネットワーク装置1のPAUSE送信時と、受信側のサーバ装置3のPAUSE開始時の2回のタイミングにおいて、それぞれPAUSE送信、またはPAUSE開始が遅れてしまう。この遅れは、ネットワーク装置1またはサーバ装置3が、データフレーム送信中である場合に生じる。この遅れの長さは、送信中のフレームの長さに対応するものである。また、サーバ装置3が、PAUSEしている期間中に、次のPAUSEフレームを受信した場合は、そのPAUSEフレームを受信した時点でPAUSEタイマ8を上書きする。

【0028】

数1の関係式がサイズの大きいデータフレーム送信の影響を受けることのないように、PAUSE送信またはPAUSE開始の遅れを考慮しなければならない。従って、サーバ装置3が送信する最大フレームサイズM（単位:バイト）に対して、PAUSE時間aおよびPAUSE送信間隔bは、数2の関係式を満たす必要がある。

【0029】

【数2】

$$b - a > 2 \times M / 64 \quad \dots \text{ (数2) }$$

ここで、64の単位は、バイト/時間であり、上述の如く、単位時間に送信されるフレームのバイト数である。

【0030】

数2について、図13を用いて説明する。図13は、PAUSE時間aとPAUSE送信間隔bとの関係を示すグラフである。201は、Mバイトのフレームを送信した場合の送信トラフィックと送信時間を示す。PAUSEフレーム送信時にネットワーク装置の送信リンク2にMバイトのフレームが送信中である場合、最大遅延時間として、 $M/64$ が発生する。なお、 $M/64$ は、Mバイトのフレームの送信に費やされる時間である。さらに、サーバ装置3でPAUSE受信後、PAUSE開始のタイミングで、サーバ装置3からの送信リンク2上でMバイトのフレームが送信中である場合、最大遅延時間として、 $M/64$ がさらに追

加される。数 1 で説明したように、P A U S E 時間 a 、 b の比率で帯域制御を行うので、P A U S E 時間 a は P A U S E 送信間隔 b の時間内に収まる必要がある。そのために、数 2 の条件式が求められる。

【0031】

さらに、サーバ装置 3 が P A U S E している期間中においては、サーバ装置 3 が搭載するアプリケーションからの送信データが、データメモリ 15 に滞留される。しかし、サーバ装置 3 のアプリケーションは、順次、送信を要求する。よって、P A U S E 期間中には、データメモリ 15 がオーバーフローすることを防止しなければならない。従って、データメモリ 15 の容量 C と P A U S E 時間 a は、数 3 の関係式を満たす必要がある。

【0032】

【数 3】

$$C > 64 \times a \quad (\text{単位: バイト}) \quad \dots (\text{数 3})$$

ここで、64 の単位は、バイト / P A U S E 時間である。

【0033】

数 3 の条件から P A U S E 時間 a の最大値が決まる。次に、数 2 から P A U S E 送信間隔 b の最小値が決まる。この P A U S E 時間 a 、P A U S E 送信間隔 b の値と、数 1 から、帯域使用率 n の値の最小値が得られる。P A U S E 送信間隔 b の値を大きく取ると、サーバの送信トラフィックの帯域使用率 n の最小値が大きくなり、ネットワーク管理者が設定できる帯域の範囲が小さくなる。以上を考慮して P A U S E 送信間隔 b の値を設定する。

【0034】

図 2 に、帯域制御プログラム 10 の処理フローを示す。

【0035】

帯域制御プログラム 10 は、ネットワーク管理者からのコマンド入力でスタートする (21)。帯域使用率 n の値は、コマンド入力時に設定するか、または事前に設定済みのパラメータ初期設定用のファイル等から読み込む。P A U S E 送信間隔 b は、予め帯域制御プログラム 10 に組込まれている。プログラム開始時に P A U S E 時間 a の値を計算し、P A U S E フレームを作成してデータメモリ

14へ保管する(22)。次に、タイマ12にPAUSE送信間隔bの値をセットする(23)。タイマ12は、512ビット時間(100Mbpsのリンクでは5.12マイクロ秒)毎に1ずつ減算される(24)。タイムアップしたとき(25)に、帯域制御プログラム10は、MAC制御部6に対して、PAUSEフレームの送信を指示(以下、PAUSE送信指示という)を発行する(26)。

【0036】

PAUSE送信指示を受けたMAC制御部6は、MACチップセット4に対し、データメモリ14に対するDMA(Direct Memory Access)を指示する。指示を受けたMACチップセット4は、PAUSEフレームの送信を実行する。DMAは、MACチップセット4内のDMA制御部により制御される。

【0037】

以上で説明した図2の一連の処理が、PAUSE送信間隔bをサイクルタイムとして繰り返される。PAUSEフレームを受信したサーバ装置3のMACチップセット5及びMAC制御部7は、指示されたPAUSE時間aの期間中、PAUSEを実行する。この実行の結果、サーバ装置3からの送信トラフィックを、リンク2の送信帯域のn%以下とすることが可能となる。

【0038】

本実施例は、図1に示すネットワーク装置1とサーバ装置3との接続のみには限られない。

【0039】

図15に示すネットワーク装置151は、図1に示すサーバ装置3と同様のネットワークインタフェース制御部を備えている。従って、サーバ装置3をネットワーク装置151と置き換えた場合にも、本実施例が適用可能である。

【0040】

また、ネットワーク装置1同士を接続させた場合にも、本実施例が適用可能である。

【0041】

なお、ネットワーク装置とは、ハブ、LANスイッチ、ルータ等の装置のみに限られず、ネットワークシステムを介して情報のやり取りを行う装置、例えばPC (Personal Computer)、サーバ等の全般をいう。

【0042】

本実施例を用いた帯域制御の様子を図9に示す。点線100は、サーバ装置3のアプリケーションが要求する送信トラフィックを示す。実線101は、本実施例を用いた帯域制御されたネットワーク装置1-サーバ装置3間のリンク2上の送信トラフィックを示す。この図では、リンク2の総帯域100Mbpsに対して、ネットワーク管理者の設定した帯域使用率 n を50%と設定した場合である。帯域使用率50%を超える分の要求トラフィックが時間的に後回しにされ、要求トラフィックの小さな時間帯へシフトしている様子が示されている。なお、設定した帯域使用率 n に対応する帯域は、50Mbps (104) である。

【0043】

なお、図9のグラフにおいてリンク2上の送信トラフィック101は常に設定帯域104以下であるように示されている。しかし、これはマクロな時間軸で表現した場合を示す。

【0044】

図10は、図9に示す時間帯102を拡大したグラフである。ミクロな時間軸で表現した場合、トラフィックが0の時間帯と、トラフィックが100Mbpsの時間帯とは交互に表現される。このうち、トラフィックが0の時間帯は、サーバ装置3がPAUSEしている状態が存在することを示す。トラフィックが100Mbpsの時間帯は、サーバ装置3によるフレームが送信されている状態が存在することを示す。時間帯102の合計トラフィックを平均した分のトラフィックは、帯域使用率 n に対応した帯域50Mbps (104) 以下となる。

【0045】

なお、本実施例は、サーバ装置3が送信するトラフィックに無関係に定期的にPAUSEフレームを送信するものである。従って、サーバ装置3が送信するトラフィックが、ネットワーク管理者により設定された帯域使用率 n 以下であって

も P A U S E 機能によりサーバ装置 3 からのデータフレームの送信が中断される。これに関して、図 9 に示す時間帯 1 0 3 の時間軸を拡大して表したものが、図 1 1 のグラフである。点線 1 0 0 は、上述のように、サーバ装置 3 のアプリケーションが要求する送信トラフィックである。図中、ミクロに表現した場合、リンク 2 上の送信トラフィックが 0 である時間帯が存在する。しかし、リンク 2 上の送信トラフィックが 0 の時間帯においても、サーバ装置 3 のアプリケーションはフレームの送信を要求する。従って、送信トラフィックが 0 の時間帯に対応したトラフィック 1 0 0 の部分は、サーバ装置 3 の送信再開時に、その時間帯に本来送信されるべきトラフィック 1 0 0 の部分と合わせて送信される。リンク 2 上の送信トラフィックが 0 の時間帯に、送信データは、一時的に止められて、データメモリ 1 5 に滞留している。送信データは、送信再開時に、一気に送出される。この場合、P A U S E 送信間隔 b の時間での平均送信トラフィック量は、影響を受けない。しかし、個々のフレーム送信に対しては、P A U S E 時間 a の値を最大値とする送信遅延が発生する。

【 0 0 4 6 】

この場合の送信トラフィック 1 0 5 の波形を、図 1 2 に正確に表現する。

【 0 0 4 7 】

次に、この送信遅延を起こさずに帯域制御を行うことが可能な実施例について説明する。

【 0 0 4 8 】

(実施例 2)

本発明の他の実施例を説明する。

【 0 0 4 9 】

図 3 のネットワーク装置 3 1 は、図 1 のネットワーク装置 1 に、新たにバイトカウンタ 3 2 を設けている。バイトカウンタ 3 2 は、ハードウェアおよびソフトウェアで構成されている。

【 0 0 5 0 】

図 3 において、バイトカウンタ 3 2 は、サーバ装置 3 からの送信トラフィックを実測するものである。帯域制御プログラム 3 3 は、バイトカウンタ 3 2 の測定

値を利用して、P A U S E 時間 e を計算する。計算された P A U S E 時間 e は実施例 1 と同様に、サーバ装置 3 に指示される。なお、本実施例においては、実施例 1 と異なり、P A U S E 時間を e で示す。

【0051】

本実施例では、バイトカウンタ 32 により実際のサーバ装置 3 の送信トラフィックを観測できる。観測をもとに、ネットワーク装置 31 は、サーバ装置 3 の送信トラフィックが、帯域使用率 n を超えたときだけ P A U S E フレームを送信する。従って、P A U S E フレームを送信するタイミングを可變的に決定するものである。

【0052】

以下、具体的に説明する。

【0053】

本発明による帯域制御プログラム 33 は、タイマ 34 とレジスタ群 35 とを有する。レジスタ群 35 は、レジスタ n 、レジスタ c 、レジスタ p を有する。

【0054】

ネットワーク管理者は、リンク 2 の帯域に対して、サーバ装置 3 からのフレームの送信に要する帯域使用率 n (%) の値を設定する。帯域使用率 n は、レジスタ群のレジスタ n にセットされる。

【0055】

帯域制御プログラム 33 は、トラフィック観測時間 c 毎に、リンク 2 のトラフィックを観測する。ネットワーク管理者は、予め、トラフィック観測時間 c を設定する。このトラフィック観測時間 c が、レジスタ c に設定される。なお、トラフィック観測時間 c とは、帯域制御プログラム 33 がバイトカウンタ 32 の測定する値を読み取る時間の間隔のことである。

【0056】

許可トラフィック p (単位: バイト) は、帯域使用率 n を基に計算される。許可トラフィック p は、レジスタ p に設定される。

【0057】

バイトカウンタ 32 は、サーバ装置 3 からの送信トラフィック (以下、サーバ

送信トラフィックという) x (単位: バイト) を測定する。帯域制御プログラム 33 は、トラフィック観測時間 c において、サーバ送信トラフィック x が許可トラフィック p を超えた場合に、PAUSE フレームを生成する。帯域制御プログラム 33 は、数 4 の関係式を満たす PAUSE 時間 e を計算する。PAUSE 時間 e を計算した後、PAUSE フレームを生成して、発行する。PAUSE 時間 e とは、実施例 1 と同様に、PAUSE フレームにより、サーバ装置 3 からのフレームの送信が止められる時間のことである。

【0058】

【数 4】

$$x / (64 \times (c + e)) \leq n / 100 \quad \dots (\text{数 4})$$

すなわち、PAUSE 機能により、トラフィック観測時間 c と PAUSE 時間 e を合計した時間内でのサーバ送信トラフィック x の帯域使用率が、ユーザにより設定される $n / 100$ 以下になるように PAUSE 時間 e を求める。

【0059】

許可トラフィック p は、PAUSE フレームの発行の判定基準となる。許可トラフィック p は、数 5 より得られる。

【0060】

【数 5】

$$p = 64 \times c \times n / 100 \quad \dots (\text{数 5})$$

ここで、64 の単位は、バイト / 時間であり、単位時間に送出されるフレームのバイト数である。

【0061】

数 4 および数 5 から、数 6 が得られる。数 6 は、PAUSE 時間 e の計算式である。実際、PAUSE 時間 e は、 $x - p > 0$ の場合に、数 6 を満たす最小の整数値となる。

【0062】

【数 6】

$$e \geq ((x - p) / 64) \times (100 / n) \quad \dots (\text{数 6})$$

トラフィック観測時間 c に対しては、図 1 の実施例と同様に、実装上の制約が

ある。ここでは、サーバ装置 3 が送信する最大フレームサイズ M 、またはサーバ装置内蔵データメモリ 15 の容量 C による実装上の制約である。この実装上の制約について、数 7 および数 8 で示す。これらの式は、それぞれ、図 1 の実施例の数 2 および数 3 に対応するものである。

【0063】

【数 7】

$$c > 2 \times M / 64 \quad \dots (\text{数 7})$$

【0064】

【数 8】

$$C > 64 \times c \times (100 - n) / n \quad \dots (\text{数 8})$$

上記のトラフィック観測時間 c 、サーバ送信トラフィック x 、許可トラフィック p 、PAUSE 時間 e の関係を図 14 に示す。トラフィック 202 は、トラフィック観測時間 c に測定されたサーバ送信トラフィック x である。トラフィック 203 は、トラフィック観測時間 c の経過の後に続けて観測される観測時間 $c + e$ の期間のうち、サーバ装置 3 により PAUSE が実行されるまでの期間のサーバ送信トラフィックである。トラフィック 204 は、トラフィック観測時間 c の経過の後に続けて観測される観測時間 $(c + e)$ の期間のうち、サーバ装置 3 により PAUSE が実行された後の期間のサーバ送信トラフィックである。

【0065】

サーバ送信トラフィック x が、数 5 の許可トラフィック p を超える場合、帯域制御プログラム 33 は、数 6 から PAUSE 時間 e を計算する。計算の後、帯域制御プログラム 33 は、PAUSE フレームの送信を指示する。ネットワーク装置 31 から PAUSE フレームが送信される。その後、サーバ装置 3 にて PAUSE が実施される。

【0066】

ネットワーク装置 31 側では、バイトカウンタ 32 が、トラフィック観測時間 c の経過の後に、続けて観測時間 $c + e$ の期間、サーバ送信トラフィック x を観測する。観測の結果、帯域制御プログラム 33 は、トラフィック 203 とトラフィック 204 とを合わせたサーバ送信トラフィック x' の値を得る。このとき、

サーバ送信トラフィック x' が、再び、許可トラフィック p を超える場合がある。この場合には、帯域制御プログラム 33 は、再び、PAUSE 時間 e' を計算して、次の PAUSE フレームの送信を指示する。以後、サーバ装置 3-ネットワーク装置 31 間においては、これらの動作が繰り返される。

【0067】

帯域制御プログラム 33 の処理フローについて、図 4 を用いて説明する。

【0068】

帯域制御プログラム 33 は、ネットワーク管理者からのコマンド入力でスタートする (41)。帯域使用率 n の値は、コマンド入力時に設定するか、又は事前に設定されたコンフィグレーションファイル等から読み込む。トラフィック観測時間 c は、予めプログラム 33 に組み込まれている。プログラム開始時に許可トラフィック p の値を計算して、レジスタ p にセットする (42)。次に、タイマ 34 にトラフィック観測時間 c の値をセットして (43)、続けてバイトカウンタ 32 をリセットする (44)。タイマ 34 は、512 ビット時間 (100 Mbps のリンクでは 5.12 マイクロ秒) 毎に 1 ずつ減算される (45)。減算の結果、タイムアップしたとき (46) に、バイトカウンタ 32 より、サーバ送信トラフィック x を読み取る (47)。サーバ送信トラフィック x が、帯域使用率 n に対応する帯域量を超えているか否かの判定を行う (48)。サーバ送信トラフィック x が、帯域使用率 n に対応する帯域量以下であれば、再度タイマ 34 に、トラフィック観測時間 c の値をセット (49) して、処理 (44) からの処理を繰り返す。

【0069】

一方、処理 (48) の判定において、サーバ送信トラフィック x が帯域使用率 n に対応する帯域量を超えている場合には、PAUSE 時間 e を計算する。計算の後、PAUSE 時間 e の値をセットした PAUSE フレームの生成を指示して、MAC 制御部 6 に対して PAUSE フレームを送信するための指示を発行する (50)。PAUSE 送信指示を受けた MAC 制御部 6 は、MAC チップ 4 に対し、PAUSE フレームが保管されているメモリ 14 に対しての DMA を指示する。MAC チップ 4 は、PAUSE フレームの送信を実行する。帯域制御プログ

ラム33は、処理(50)に続けて、タイマ34に新たに観測時間($c+e$)の値をセットする(51)。セットの後、処理(44)からの処理を繰り返す。タイマ34に($c+e$)の値をセットする理由を述べる。PAUSE送信等がデータフレーム送信中に発行される。この場合に、PAUSE送信指示から実際のPAUSE開始までにタイムラグが生じる。従って、サーバ装置3が送信可能な時間の延べ時間がトラフィック観測時間 c になるように観測する必要があるからである。

【0070】

図4を用いて説明した一連の処理が繰り返されることにより、指示されたPAUSE時間 e の間、PAUSEが実施される。これにより、サーバ送信トラフィック x を、リンク2の帯域使用率 n 以下とすることが可能となる。さらに、PAUSEフレームの送信は、サーバ送信トラフィック x が帯域使用率 n を超えたときのみ実行される。従って、帯域使用率 n 以下の間は、サーバ送信トラフィック x による不要な送信遅延を生じることがない。

【0071】

以上、本発明の帯域制御機構について説明した。

【0072】

(実施例3)

次に、本発明の帯域制御機構を、拡張したネットワークシステムに適用した実施例について説明する。

【0073】

本実施例に示すネットワークシステムを、図5に示す。このネットワークシステムにおいて、ネットワーク装置は、複数の物理リンクを収容する。複数のサーバ装置は、複数の物理リンクの各々に接続される。

【0074】

本実施例では、ネットワーク装置に、LANスイッチを用いて説明する。LANスイッチは、本発明の帯域制御機構を備えているものとする。但し、本実施例は、LANスイッチに限らず、IEEE802.3x規格の技術をサポートしている全てのネットワーク装置に適用可能である。

【0075】

LANスイッチA(61)は、サーバ装置a(62)からサーバ装置d(65)の4台により構成されるサーバ装置(以下、「サーバ装置群」という)に接続されている。接続には、全二重通信を行えるリンクが用いられる。LANスイッチA(61)は、4台のサーバ装置のそれぞれと一台毎に接続できるポート(図示せず)を4個備えている。LANスイッチA(61)は、実施例1で示した帯域制御プログラム10、または実施例2で示した帯域制御プログラム33(以下、「帯域制御プログラム10、33」という。)を有する。LANスイッチA(61)の個々のポートには、それぞれ、MACチップおよびMAC制御部が設けられる。それぞれのMACチップおよびMAC制御部は、帯域制御プログラム10、33により制御される。帯域制御プログラム10、33の個数に対する限定はなく、個々のポートに対してそれぞれ設けられるものでもよい。また、全てのポートに対して、ただ一つ設けられるものでもよい。

【0076】

LANスイッチA(61)は、サーバ装置群と、クライアント群との間の通信を中継する位置にある。クライアント群は、LANスイッチ1(66)およびLANスイッチ2(67)に収容される。LANスイッチ1(66)およびLANスイッチ2(67)は、LANスイッチA(61)とリンクを介して接続する。

【0077】

サーバ装置a(62)からサーバ装置d(65)のそれぞれは、それぞれ別個のアプリケーションを搭載している。各々のサーバ送信トラフィックを、それぞれ、サーバ送信トラフィックa、b、c、dとする。ネットワークシステムのネットワーク管理者は、LANスイッチA(61)において、サーバ送信トラフィックaからd(68)のそれぞれに対して帯域使用率 n_a 、 n_b 、 n_c 、 n_d (図5の矢印aからdの幅が使用帯域を示す)を設定する。帯域制御プログラム10、33を起動することにより、サーバ装置群からクライアント群へ流れるサーバ送信トラフィックに対して、アプリケーション毎に所望の帯域を設定することが可能となる。

【0078】

また、本発明の帯域制御機構は、ネットワーク装置と1台のサーバ装置とを接続するリンクが、複数の場合にも適用可能である。この場合、各々のリンクに接続するポートに対応して、それぞれMAC制御部が設けられる。それぞれのMAC制御部を独立に動作させることにより、個々のリンク単位にサーバ送信トラフィックを制御することが可能である。1台のサーバ装置が複数のアプリケーションを搭載する場合に有効である。

【0079】

現在、IEEE 802.3adでは、リンク・アグリゲーション(LA)規格についての標準化作業が進行中である。図8は、リンク・アグリゲーション(LA)規格を適用した場合のOSI(Basic Reference Model)基本参照モデルのデータリンク層におけるプロトコルの関係を示している。以下、図8を用いて説明する。

【0080】

本発明の帯域制御プログラム10、33の実装方法には、以下の2通りがある。1通り目は、帯域制御プログラム10、33を、LAサブレイヤ92の上位プログラム93として実装するものである。この場合、帯域制御プログラム10、33は、図8に示すMACクライアント93に相当する部分に該当する。LAサブレイヤ92は、1つ存在して、複数のMAC制御部91を集約する。LAサブレイヤ92は、複数のMAC制御部91を、上位のプログラムであるMACクライアント93に対して、1個のMACインタフェースとして見せる。従って、帯域制御プログラム10、33がPAUSE送信指示を発行した場合、LAサブレイヤ92は、PAUSE送信指示を、下位の全てのMAC制御部91に配布する。特に、実施例2の帯域制御機構の場合、帯域制御プログラム33は、バイトカウンタ32を使用する。バイトカウンタ32は、LAサブレイヤ92の配下にある全てのリンクのトラフィックを集計する。従って、バイトカウンタ32は、帯域制御プログラム33に集計データを伝えるように実装される。

【0081】

この1通り目の実装方法が適用されるネットワークシステムの一例を挙げる。

この一例には、ネットワーク装置が、1 台のサーバ装置につき、複数のリンクで接続される場合が該当する。サーバ装置とネットワーク装置とは、それぞれ、複数のポートを介して複数のリンクに接続される。この場合、それぞれのポートに対応するMAC制御部 9 1 は、上位の制御プログラムであるMACクライアント 9 3 で 1 つにまとめられて制御される。

【0 0 8 2】

この場合のネットワークシステムの構成を、図 6 に示す。

【0 0 8 3】

図 6 において、7 2 から 7 5 の各サーバ装置は、複数のリンクでLANスイッチ A (7 1) と接続している。複数のリンクは、LA で束ねられたものである。図 5 の場合と同様に、各サーバ装置 (7 2 から 7 5) は、それぞれ別個のアプリケーションを搭載している。各々のサーバ送信トラフィックは、サーバ送信トラフィック a、b、c、d とする。LANスイッチ A (7 1) は、本発明の帯域制御機構を有する。このネットワークシステムのネットワーク管理者は、LANスイッチ A (7 1) において、帯域使用率 n を設定する。帯域使用率 n は、サーバ送信トラフィック a から d (7 8) に対して、それぞれ、帯域使用率 n_a 、 n_b 、 n_c 、 n_d と設定するものとする。帯域使用率 n の設定後、帯域制御プログラム 1 0、3 3 を起動する。これにより、サーバ装置群からクライアント群へ流れるサーバ送信トラフィックに対して、アプリケーション毎に所望の帯域を設定することが可能となる。

【0 0 8 4】

具体的には、図 1 に示す実施例 1 の帯域制御プログラム 1 0 では、各々のリンクに対して定期的にPAUSEフレームを送信する。従って、送信間隔 b および PAUSE 時間 a は、個々のリンクの帯域をもとに個別に計算する。一方、図 3 に示す実施例 2 の帯域制御プログラム 3 3 では、LA で束ねられた全てのリンクの帯域の総和に対して帯域使用率 n を設定する。トラフィック観測時間 c および PAUSE 時間 e の計算も全てのリンクのトラフィックの集計値をもとに計算する。

【0085】

2通り目は、帯域制御プログラム10、33は、LAサブレイヤ92と並列に、個々のMAC制御部91に直接アクセスできるように実装するものである。この場合、帯域制御プログラム10、33は、図8に示すMACクライアントプログラム94に相当するものである。この場合、帯域制御プログラム10、33は、LAサブレイヤ92の存在を無視して、個々のリンク単位で独立に動作する。なお、現時点ではIEEE802.3においてこの位置にMACクライアントを置くことは規定されていない。

【0086】

この場合のネットワークシステムの構成の一例を、図7に示す。図7においては、図6と同様に、各サーバ装置(82、83)は、複数のリンクでLANスイッチA(81)と接続している。しかし、図6とは、次の点で異なる。1つは、各サーバ装置(82、83)が、各々が複数のアプリケーションを搭載している点である。もう1つは、アプリケーション毎に異なるリンク(84、85、86、87)を使用している点である。

【0087】

以上で説明したように、本発明によれば、図5、図6、図7に示すネットワークシステムにおいて、ネットワーク管理者が各サーバ装置またはリンクに対して設定した帯域使用率 n 以下になるように、送信トラフィックを制御することが可能である。

【0088】

なお、ネットワーク管理者が帯域使用率 n 等を設定する単位は、アプリケーション毎に限らず、例えば、宛先IPアドレス等を設定単位としてサーバ送信トラフィックを帯域制御することも好ましい。宛先IPアドレスを設定単位としてサーバ送信トラフィックを帯域制御する場合には、全てのサーバ装置が同一のアプリケーションを備えていることが条件となる。さらに、各々のサーバ装置は帯域制御の対象となるリンクに対して、宛先IPアドレスをもとに送信フレームを振り分けることも条件となる。

【 0 0 8 9 】

なお、本実施例では、ネットワーク装置を L A N スイッチとして説明したが、これがルータであるでも本発明を適用することが可能である。さらに、ルータにおいても、L A N スイッチを用いた場合と同様の上記効果が得られる。

【 0 0 9 0 】

【発明の効果】

本発明によれば、サーバクライアントシステムを構成するネットワークシステムにおいて、輻輳を回避させることが可能である。さらに、サーバ・トラフィックをネットワーク管理者が設定する単位（アプリケーションや宛先 I P アドレス等）で、帯域を制御することが可能である。

【図面の簡単な説明】

【図 1】

本発明によるネットワーク装置 1 を用いたネットワークの構成図である。

【図 2】

帯域制御プログラム 1 0 の処理フローを示すフローチャートである。

【図 3】

本発明によるネットワーク装置 3 1 を用いたネットワークの構成図である。

【図 4】

帯域制御プログラム 3 3 の処理フローを示すフローチャートである。

【図 5】

本発明が適用されるネットワークシステムの一構成例である。

【図 6】

本発明が適用されるネットワークシステムの一構成例である。

【図 7】

個々のリンク毎に帯域制御するネットワークシステムの構成例である。

【図 8】

本発明をリンク・アグリゲーションと共に適用した場合の O S I のデータリンク層におけるプロトコルの関係を示す構成図である。

【図 9】

サーバ送信トラフィックのグラフである。

【図 10】

図 9 に示す時間帯 102 の時間軸を拡大したグラフである。

【図 11】

図 9 に示す時間帯 103 の時間軸を拡大したグラフである。

【図 12】

図 11 のサーバ送信トラフィック 105 の波形を拡大したグラフである。

【図 13】

PAUSE 時間 a と PAUSE 送信間隔 b との関係を示すグラフである。

【図 14】

トラフィック観測時間 c、サーバ送信トラフィック x、許可トラフィック p、PAUSE 時間 e の関係を示すグラフである。

【図 15】

従来のネットワーク装置 151 を用いたネットワークの構成図である。

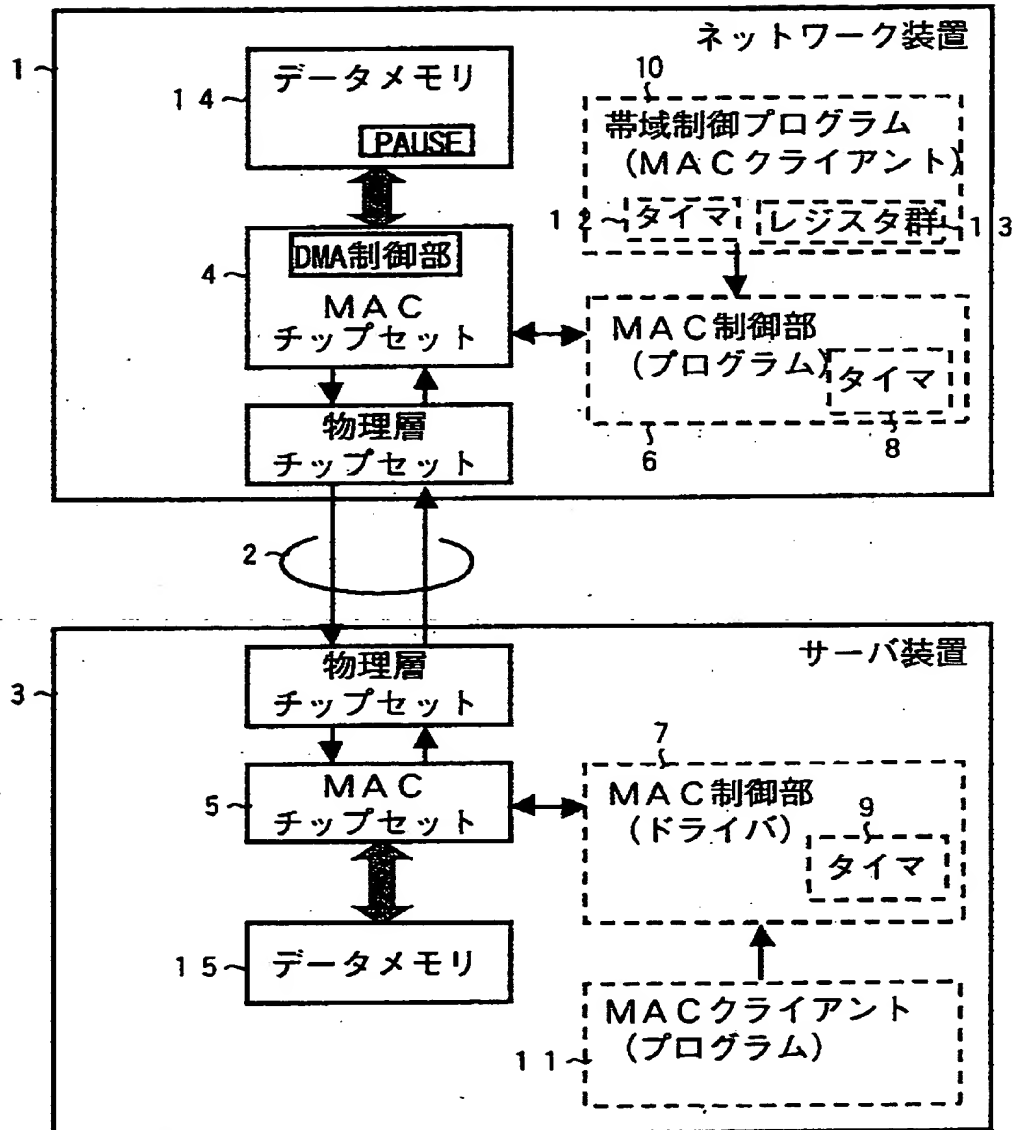
【符号の説明】

- 1、31、151・・・ネットワーク装置
- 2・・・リンク
- 3・・・サーバ装置
- 4、5・・・MACチップセット
- 6、7・・・MAC制御部
- 8、9・・・PAUSEタイマ
- 10、33・・・帯域制御プログラム
- 11・・・MACクライアントプログラム
- 12、34・・・タイマ
- 13、35・・・レジスタ群
- 32・・・バイトカウンタ

【書類名】 図面

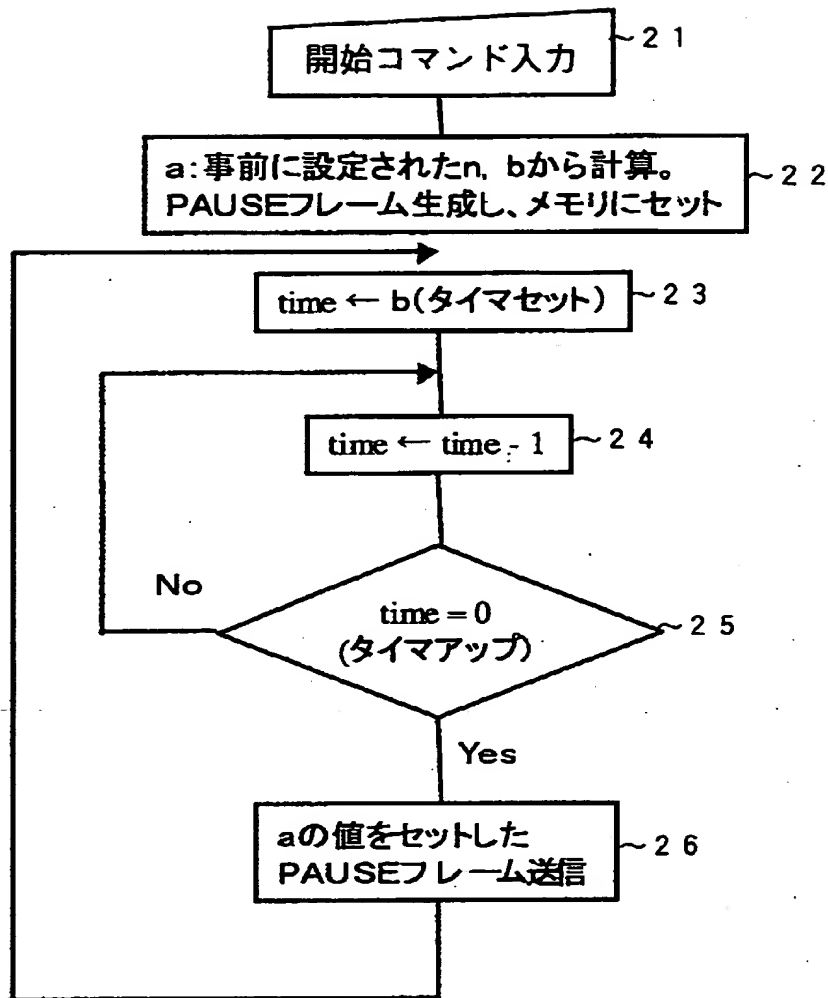
【図 1】

図 1



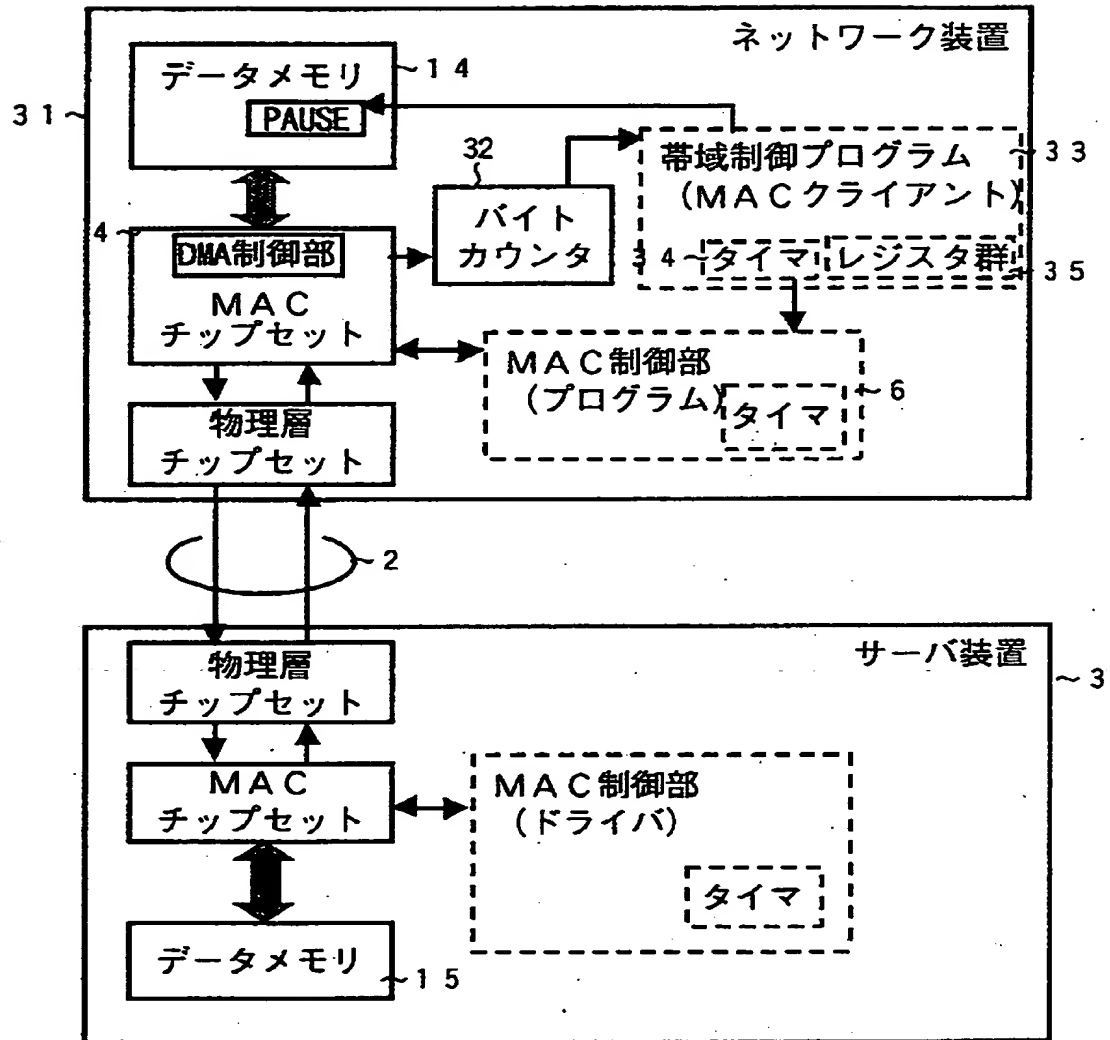
【図 2】

図 2



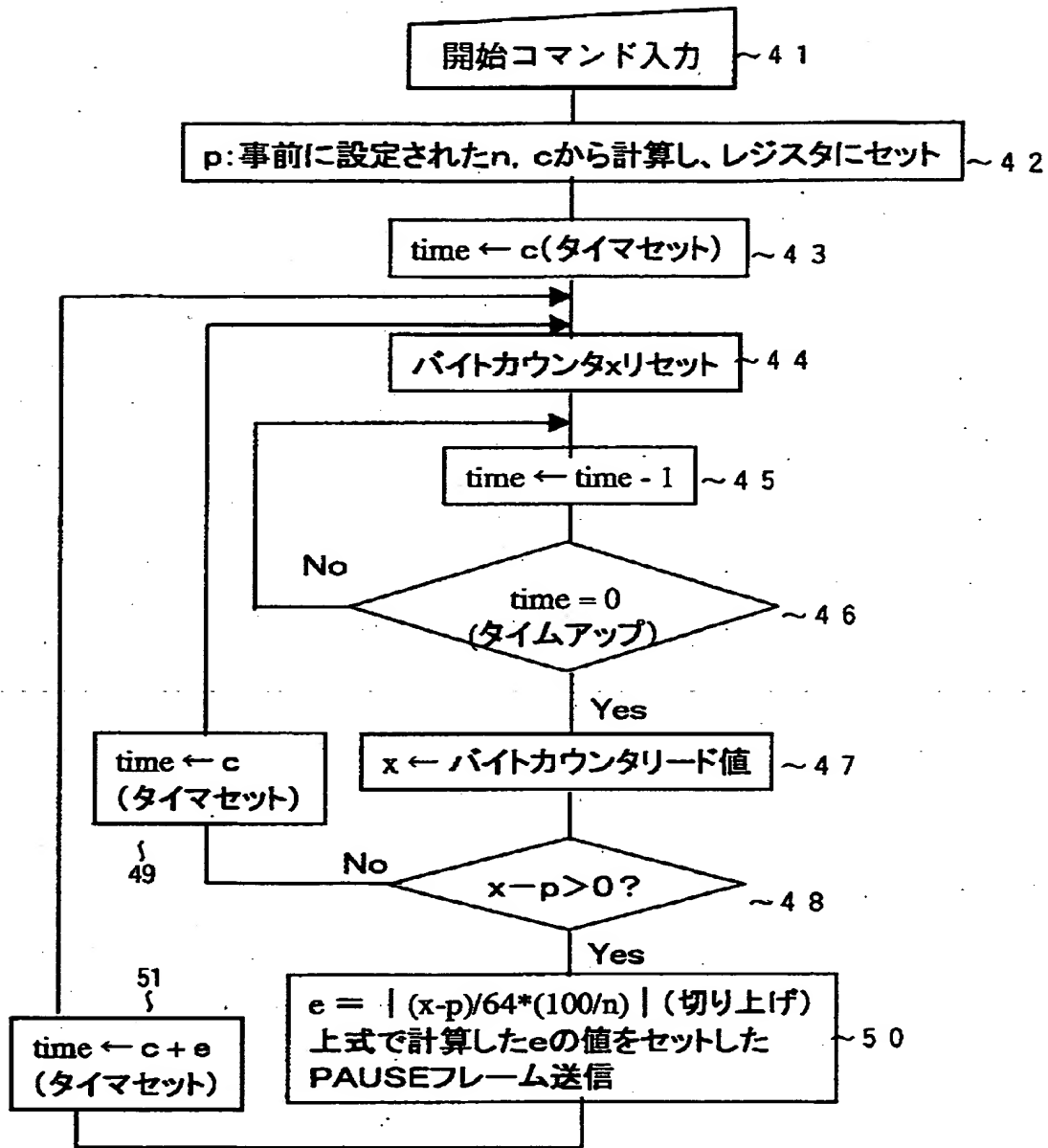
【図 3】

図 3



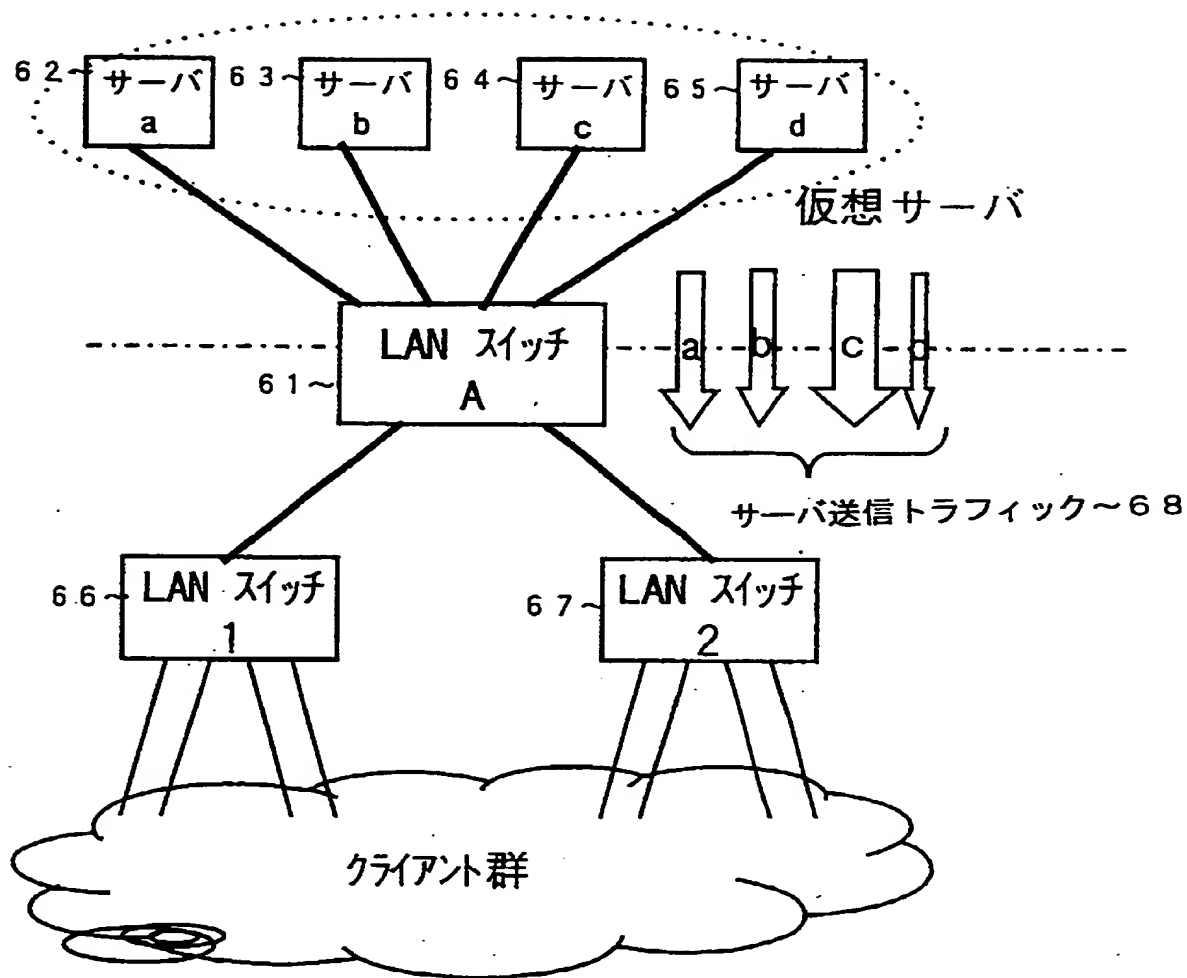
【図 4】

図 4

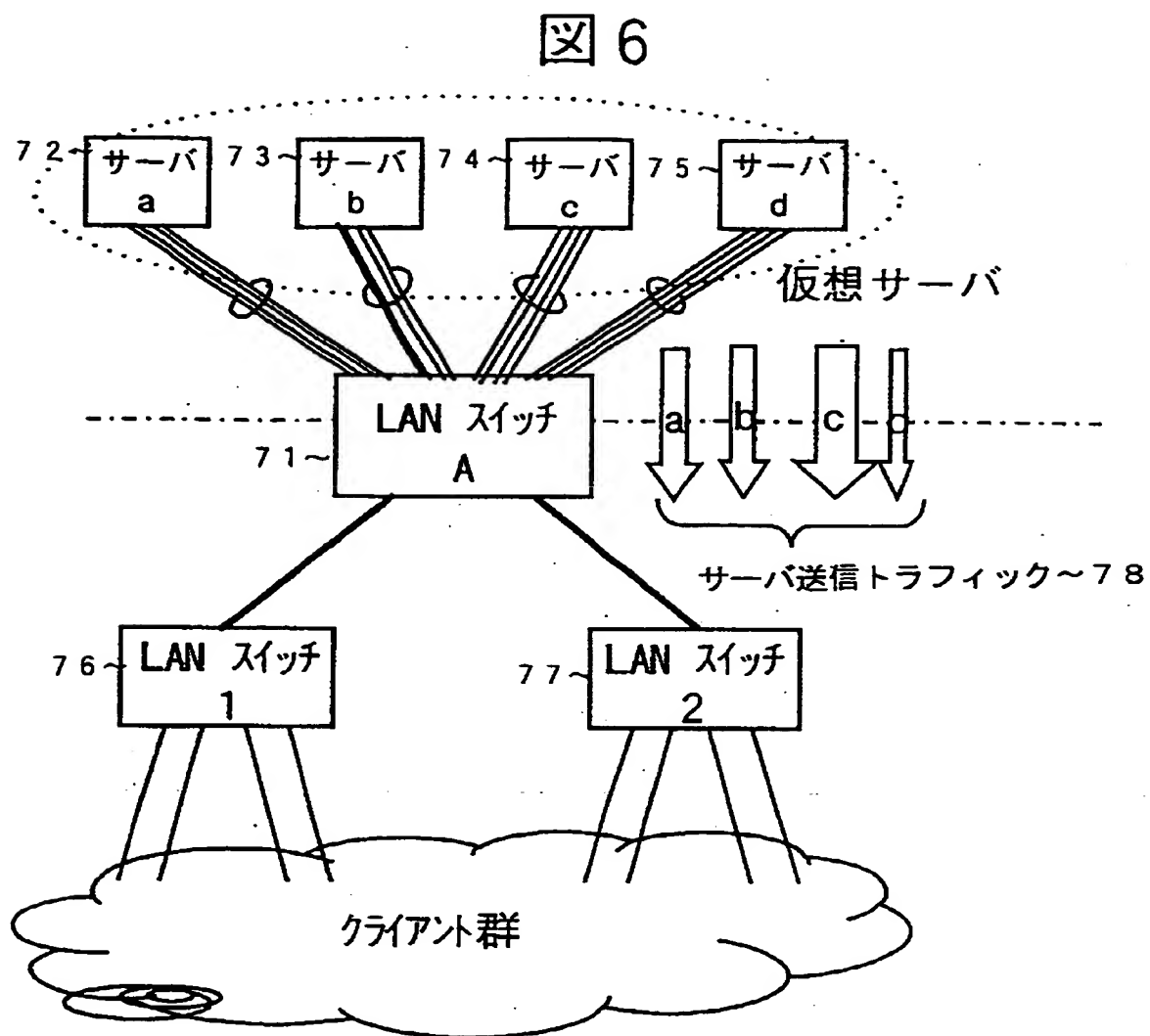


【図 5】

図 5

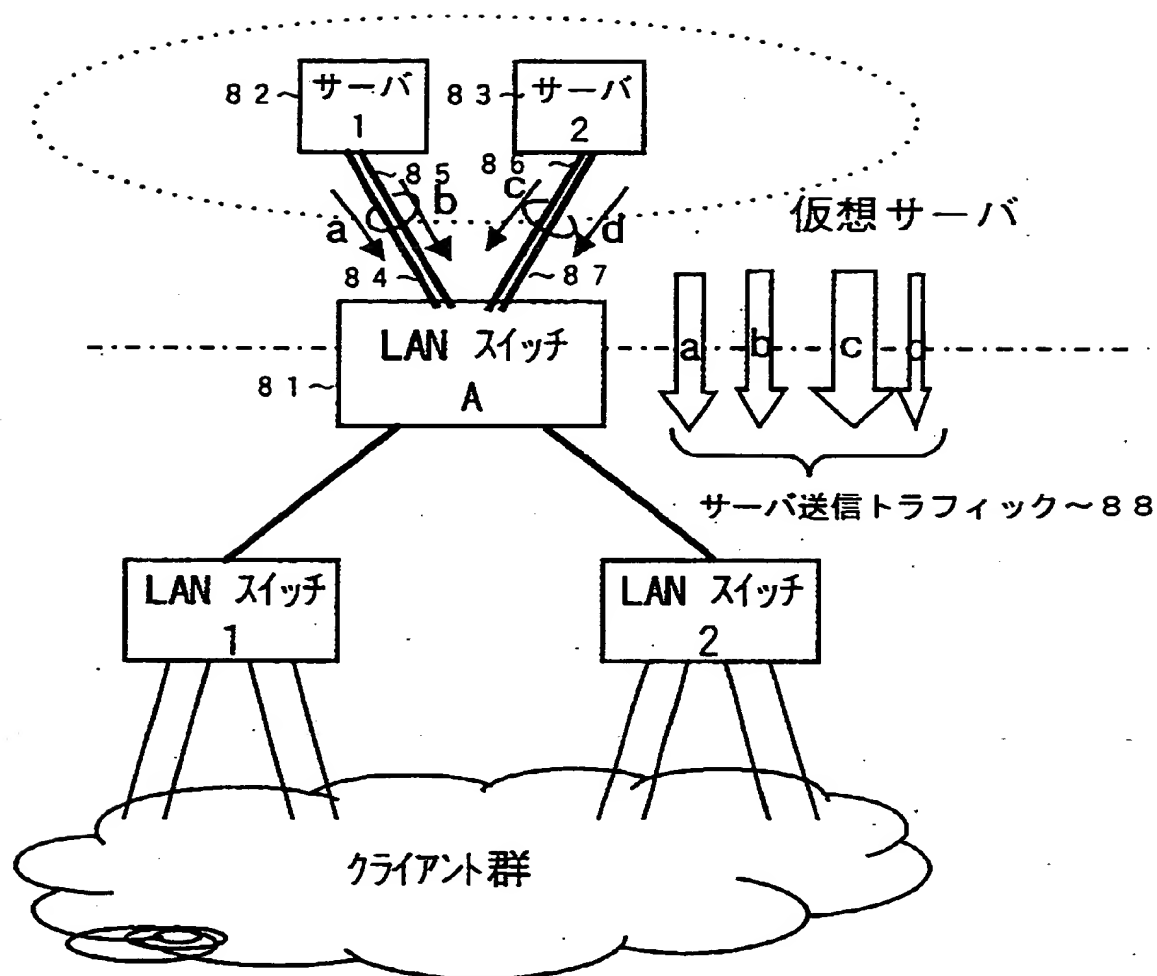


【図 6】



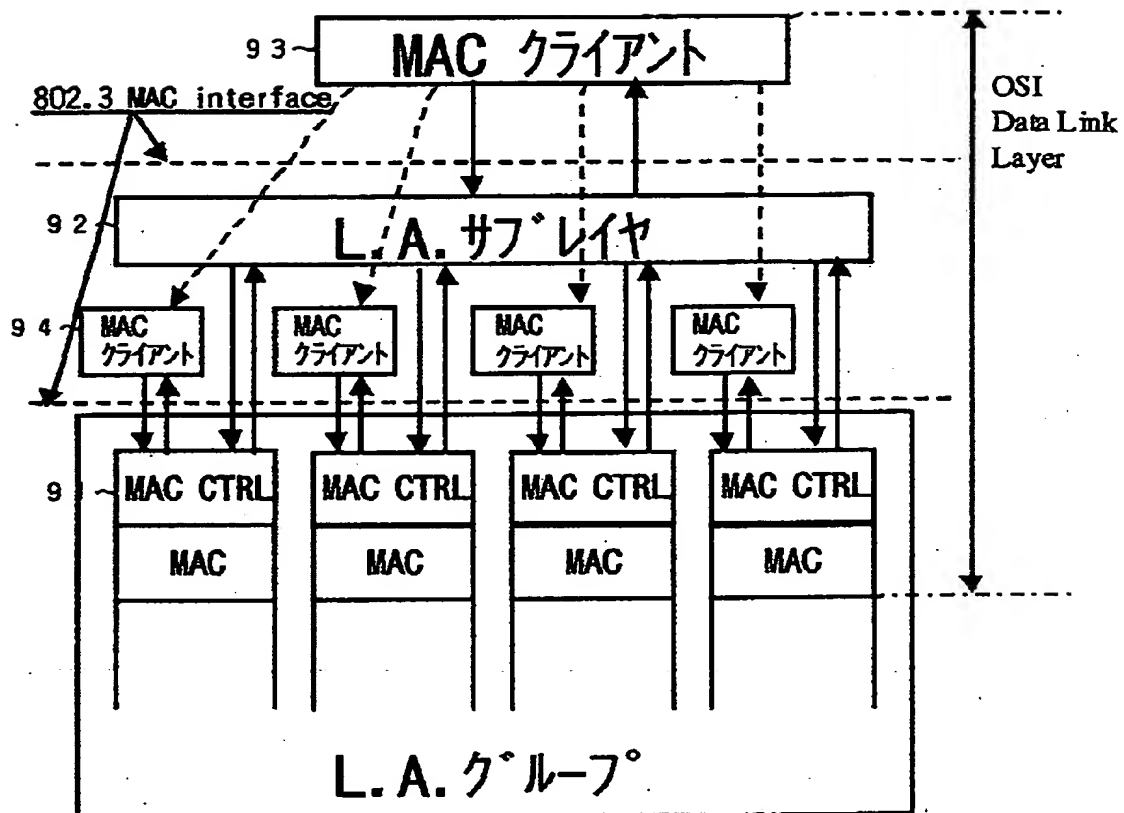
【図 7】

図 7



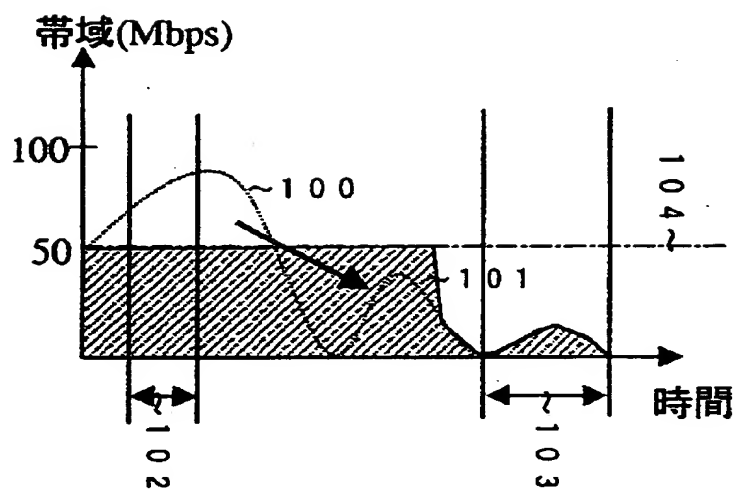
【図 8】

図 8



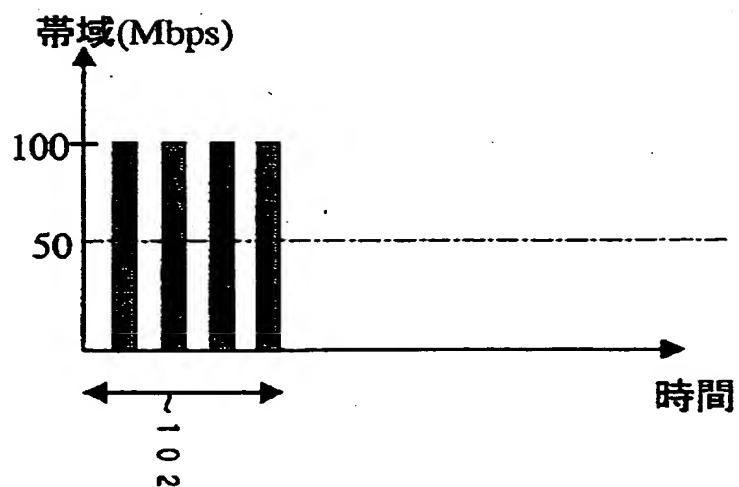
【図 9】

図 9



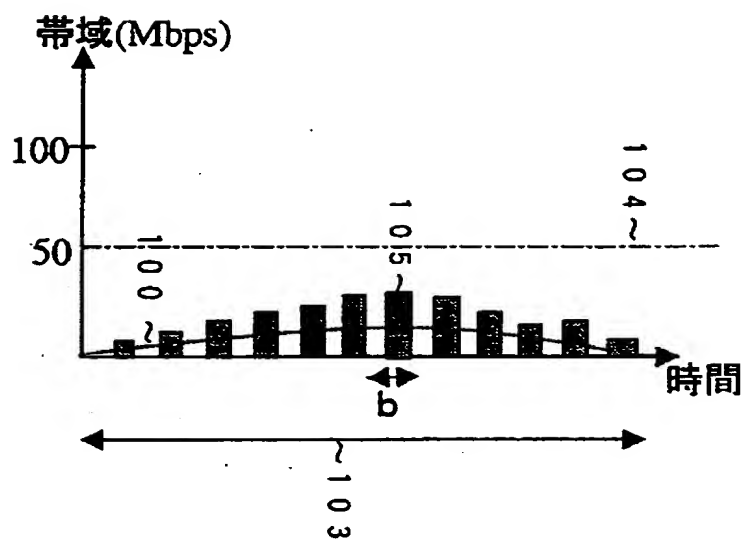
【図 1 0】

図 1 0



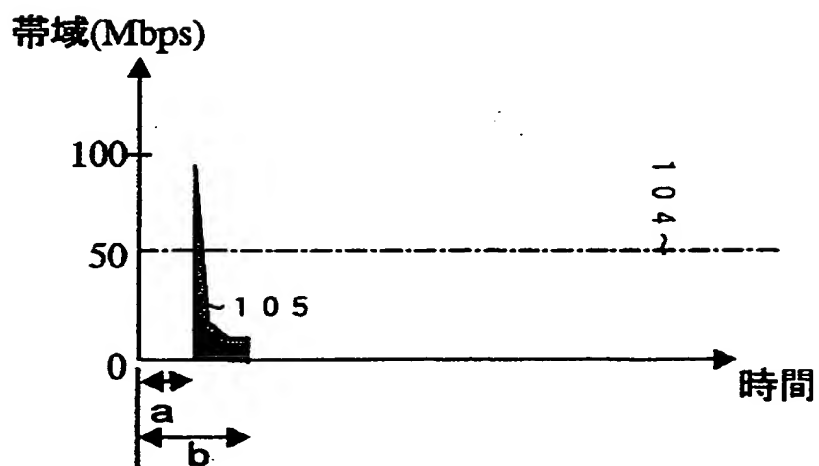
【図 1 1】

図 1 1



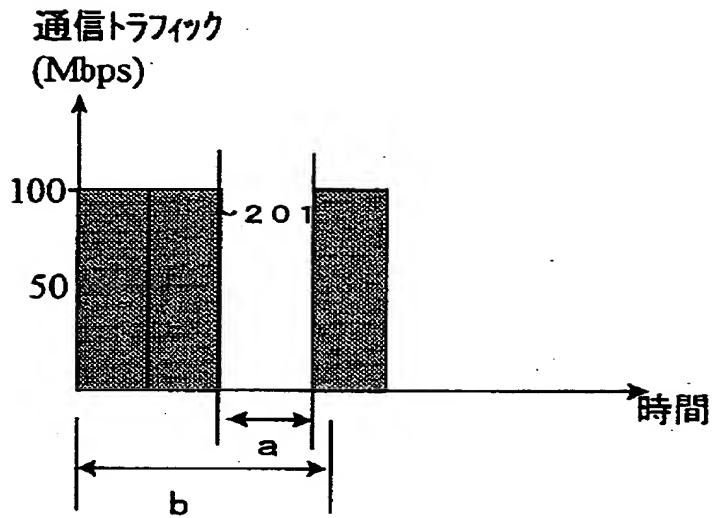
【図 1 2】

図 1 2



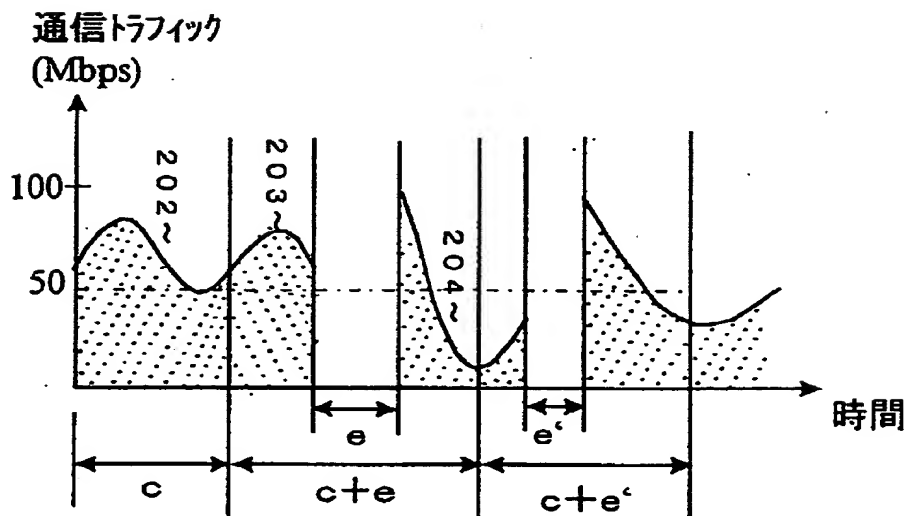
【図 1 3】

図 1 3



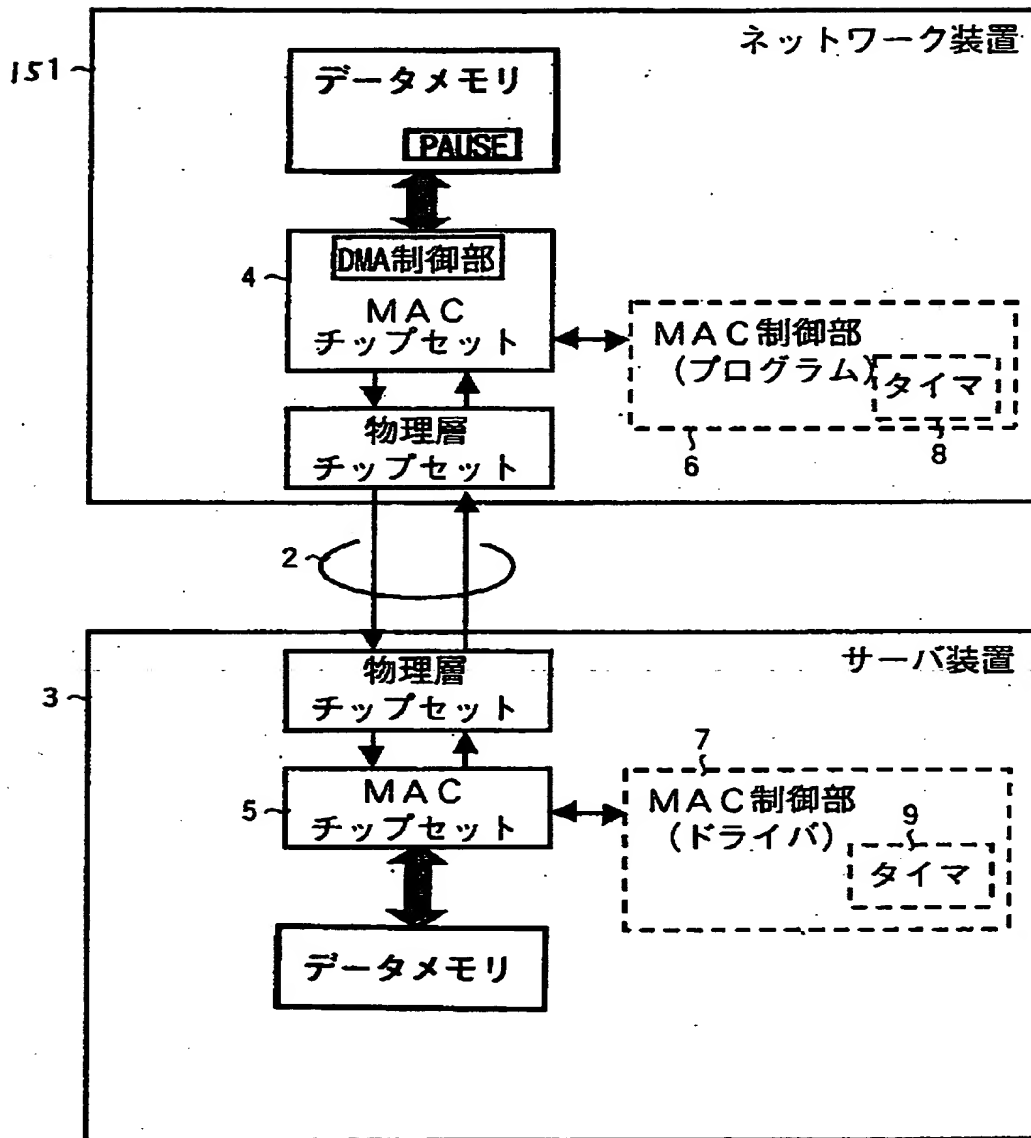
【図 1 4】

図 1 4



【図 15】

図 15



【書類名】 要約書

【要約】

【課題】

サーバクライアントシステムを構成するネットワークシステムにおいて、輻輳を回避させることが可能なネットワーク装置を提供する。

【解決手段】

ネットワーク装置 1 は、接続されたサーバ装置 3 に対して、制御フレームを送信する手段を有する。制御フレームは、指定期間のデータ送信を中断させるよう指示するものである。ネットワーク装置 1 は、制御フレームの送信を定期的を実施するよう指示する帯域制御プログラム 10 を有する。ネットワーク管理者は、予め、帯域使用率 n を設定する。これにより、サーバ装置 3 からのサーバ送信トラフィックを帯域使用率 n 以内に制御する。

【選択図】 図 1

出 願 人 履 歴 情 報

識別番号 [0 0 0 0 0 5 1 0 8]

1. 変更年月日	1 9 9 0 年 8 月 3 1 日
[変更理由]	新規登録
住 所	東京都千代田区神田駿河台 4 丁目 6 番地
氏 名	株式会社日立製作所